STUDY ON NON-PARAMETRIC EXTRAPOLATION METHOD OF TRACTOR PTO LOAD BASED ON DBSCAN

/

基于 DBSCAN 对拖拉机 PTO 的载荷非参数外推方法研究

Yin TANG[1)], Shuaijie MA[1)], Fuxi SHI[1*)], Huipeng QIU[2)], Xiao ZHANG[2)], Jianmin GAO[2)]

[1)] College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling 712100 / China

[2)] Automotive Transmission Engineering Research Institute, FAST Auto Drive Group Co., Ltd, Xi'an 710119 / China

Tel: +8613541853577; E-mail: 1135209658@qq.com

DOI: https://doi.org/10.35633/inmateh-79-04

Keywords: *Tractor, Load spectrum, Non-parametric extrapolation, DBSCAN clustering, Kernel density estimation*

ABSTRACT

The power take-off (PTO) shaft of a tractor is subjected to complex and variable random loads during field operations, and the accuracy of the resulting load spectrum directly affects the reliability of fatigue life prediction. To address the limitations of traditional parametric extrapolation methods, particularly their inadequate fitting performance, a non-parametric extrapolation method based on DBSCAN clustering is proposed. Field test validation demonstrates that the proposed method effectively captures the multi-modal distribution characteristics of the load. The extrapolated results show good agreement with the measured data in terms of cycle count distribution and pseudo-damage indicators. Compared with the fixed-bandwidth method, the proposed approach increases the coefficient of determination by 3.401% and reduces the mean squared error (MSE) by 39.154%, while yielding pseudo-damage values closer to 1. These findings provide a novel technical approach for the construction of load spectra under complex operating conditions of agricultural machinery.

摘要

拖拉机动力输出轴在田间作业过程中承受复杂多变的随机载荷，其载荷谱编制的准确性直接影响疲劳寿命预测的可信度。针对传统参数化外推方法难以精确拟合的问题，提出一种基于 DBSCAN 聚类的非参数外推方法。田间实测数据验证表明，该方法能有效刻画载荷的多模态分布特征，外推结果在循环计数分布与伪损伤等指标上与实测数据具有良好的一致性。结果表明本文提出的方法决定系数相比于固定带宽提高了 3.401%，均方误差降低了 39.154%，伪损伤值更接近 1。研究结果为农业装备复杂工况下载荷谱编制提供了新的技术途径。

INTRODUCTION

In the development of modern agriculture, tractors are widely used and play a crucial role in agricultural mechanization and efficiency improvement *(Yu et al., 2023; Yang et al., 2018)*. The tractor power take-off (PTO) shaft is an important component of the tractor transmission system and a key part affecting the durability and reliability of the entire tractor. The tractor load spectrum can reflect the load history experienced by the whole tractor or its key components, serving as the foundation for tractor fatigue analysis and reliability assessment *(Liu et al., 2011)*. Load extrapolation is a critical step in load spectrum compilation, and its effectiveness directly affects the accuracy of the compiled load spectrum. Therefore, researching load extrapolation methods suitable for the tractor PTO shaft is of great significance for improving the credibility of tractor durability and reliability tests, as well as enhancing the operational performance of the tractor *(Wang et al., 2022)*.

In recent years, the combination of clustering algorithms and kernel density estimation has provided new ideas for load extrapolation (*Chen et al., 2017*). Kernel density estimation is a statistical method that directly infers the population probability density surface based on the sample data itself without any prior theoretical distribution assumptions. Research has shown that grouping load data using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm allows for the calculation of optimal bandwidths for different clusters of data, thereby effectively addressing the limitations of a global fixed-bandwidth. The DBSCAN algorithm can identify clusters of arbitrary shapes and automatically exclude noise points, which aligns naturally with the multi-operating condition and multi-modal distribution characteristics commonly found in load data *(He et al., 2018; Wang et al., 2016)*.

Based on the above analysis, this paper proposes a non-parametric extrapolation method for tractor PTO loads based on DBSCAN clustering *(Li et al., 2025)*. First, the PTO torque time history under typical operating conditions is collected through field tests. Second, the rain-flow counting method is used to extract the mean and amplitude of load cycles, and the DBSCAN algorithm is applied to perform cluster analysis on the mean-amplitude data. The rain-flow counting method is a counting method for calculating the amplitude and mean of load cycles, which conforms to the inherent characteristics of fatigue loads and is widely used in the statistical analysis of load time history. Then, based on the clustering results, bandwidth-optimized kernel density estimation is performed for the data of each cluster to establish a load probability distribution model. Finally, extrapolation is carried out through Monte Carlo simulation to obtain a PTO load spectrum covering the full life cycle. The research results can provide more accurate load input for tractor PTO fatigue strength design and bench durability tests, while also offering a methodological reference for load spectrum compilation under complex operating conditions in the field of agricultural equipment.

MATERIALS AND METHODS

Testing System Setup

The tractor wireless torque measurement system is mainly composed of a torque sensor, a data acquisition module, a wireless transmission module, and host computer software. The wireless torque sensor, model TQ201, primarily consists of a strain measurement circuit, a data processing unit, a wireless communication unit, and a power supply unit. The strain gauges and the sensor are fixed on the output shaft, forming the basic sensor. When the output shaft rotates, the resistance of the strain bridge changes, causing the voltage at the measurement terminal to change. The voltage signal is transmitted to the data processing unit, where it passes through the node's built-in high-precision resistive bridge circuit and amplification conditioning circuit. Subsequently, the signal is transmitted to the wireless receiving module via the wireless communication unit. Throughout the entire process, the power supply unit provides power to all components.

A schematic diagram of the ultimately constructed wireless torque testing system is shown in Figure 1 below.

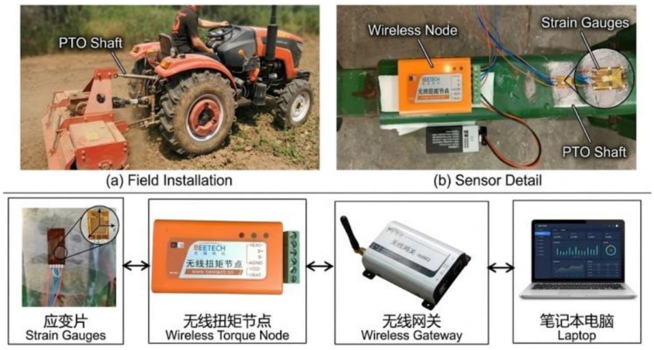

Fig. 1 – Schematic diagram of wireless torque testing system

Field Test Load Collection

From May 20 to 25, 2025, a field test of tractor power take-off (PTO) shaft torque was conducted in the suburbs of Yangling District, Xianyang City, Shaanxi Province. The test prototype was a YCX504D tractor, and the operating condition was rotary tillage using a 1GQN-180Z rotary tiller. Based on the tractor wireless torque testing system, the PTO shaft torque load is displayed and recorded in real time.

Load Signal Preprocessing

During the real-vehicle testing of tractor PTO torque, the original load signal is often contaminated with singular points that do not represent true loads due to the harsh field environment, electrical system interference, and occasional instability of the sensor. These abnormal data can severely distort the statistical characteristics of the load distribution, thereby affecting the accuracy of fatigue damage calculations. Therefore, this study adopts a strategy combining the amplitude threshold method and the gradient threshold method to eliminate and smooth the original load signal.

The amplitude threshold method is primarily used to identify and eliminate outliers that significantly exceed the extreme values of the physical system or violate statistical laws. For load signals with stationary or pseudo-stationary random characteristics, such as tractor PTO torque, this study employs the classic criterion as the basis for determination.

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{2}$$

According to statistical principles, if the data obey or approximately obey a normal distribution, the probability of abnormal points is extremely small. The discrimination criterion of the amplitude threshold method is as follows.

$$|x_i - \mu| > K\sigma \tag{3}$$

Among them, x_i is the load amplitude at the sampling point, μ is the mean of the load signal sequence, σ is the standard deviation of the load signal sequence, and K is the threshold coefficient, typically taken as 3.

When the signal point satisfies the above condition, it is determined as a gross error. To maintain the continuity of the time series, the point is usually not directly removed; instead, it is replaced by linear interpolation using adjacent valid data points, i.e.

$$x_i^* = \frac{x_{i-1}+x_{i+1}}{2} \tag{4}$$

Although the amplitude threshold method can effectively eliminate global extreme outliers, it cannot identify local abrupt signals where the amplitude does not exceed the limit but unreasonable mutations occur between adjacent sampling points. When a tractor encounters gravel or experiences sudden implement jamming, certain impacts may occur; however, limited by the inertia of the mechanical system, the instantaneous rate of change of the load has a physical upper bound. Therefore, the gradient threshold method is introduced to eliminate high-frequency abnormal sudden-change signals.

The gradient threshold method determines anomalies by calculating the difference between two adjacent sampling points (i.e., the local gradient) and comparing it with a set critical threshold. Its discrimination criterion can be expressed as:

$$|\Delta x_i| = |x_i - x_{i-1}| > \varepsilon \tag{5}$$

To avoid misidentifying actual transient impact loads as outliers, it is usually necessary to perform bidirectional gradient verification in conjunction with adjacent points. That is, a data point is determined as an outlier only when the gradients between it and both the previous and subsequent points are greater than the set threshold and are in opposite directions.

The improved discrimination criterion is as follows:

$$\begin{gathered} |x_i - x_{i-1}| > \varepsilon \\ |x_i - x_{i+1}| > \varepsilon \\ (x_i - x_{i-1})(x_i - x_{i+1}) > 0 \end{gathered} \tag{6}$$

In the formula, x_i is the current load signal point to be detected; x_{i-1} is the previous valid load signal point; x_{i+1} is the next load signal point; and ε is the gradient threshold. This threshold is typically determined based on the physical characteristics (maximum stiffness and response frequency) of the tractor's power transmission system or suspension system.

When a data point is identified as a gradient anomaly, polynomial interpolation is similarly employed for smoothing correction to restore the true physical profile of the signal, ensuring the accuracy of cycle amplitude extraction during subsequent rain-flow counting.

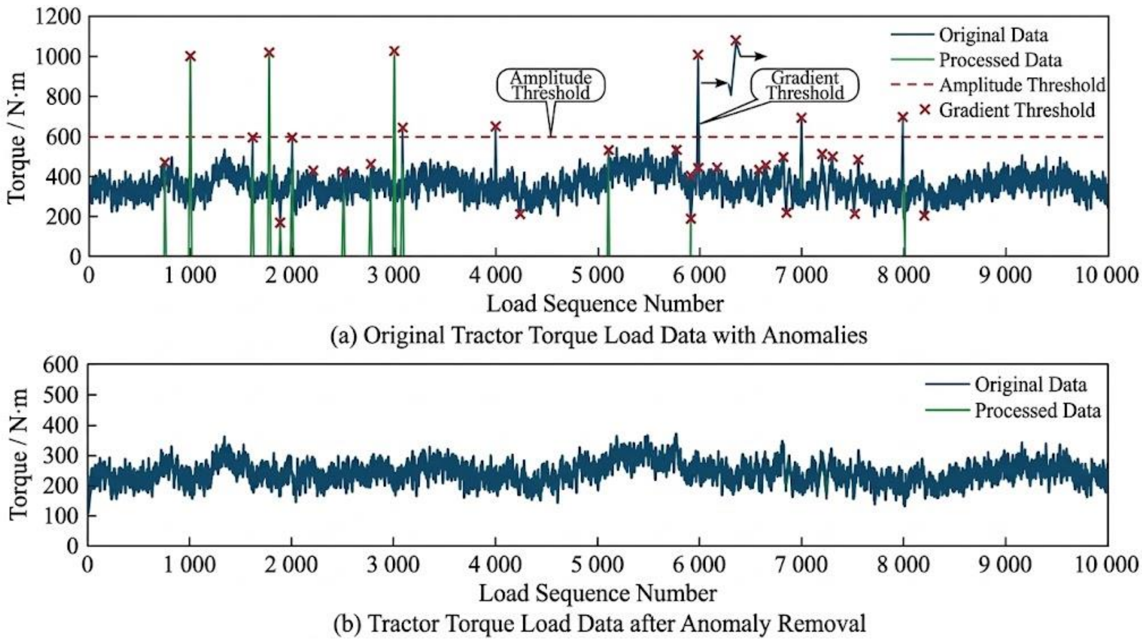

Fig. 2 – Comparison chart for removing singular points

After removing singular points from the load signal, the signal often still contains random high-frequency interference caused by tractor engine vibration, high-frequency excitation from the transmission system, and the inherent thermal noise of the sensor. To ensure that the subsequent rain-flow counting method can accurately extract the true load amplitudes that affect fatigue life, further smoothing and noise reduction of the signal are required.

For load spectrum compilation, traditional moving average filtering tends to flatten the peaks and valleys of the signal, thereby underestimating the true fatigue damage. Therefore, this paper adopts the S-G polynomial smoothing filter method for signal noise reduction. The core advantage of the S-G filter lies in its ability to filter out high-frequency noise while preserving the local extreme characteristics and waveform width of the original load signal to the greatest extent.

The S-G filter is essentially a sliding window filtering method based on local polynomial least squares fitting. Let the width of the smoothing sliding window be 2m+1; the set of data points within the window is denoted as $\{(ti,xi)|i=-m, ..., 0, ..., m\}$, where the center point corresponds to i=0. Within each sliding window, a k-order polynomial is used to fit the data points in the window.

The fitting polynomial can be expressed as:

$$p(j) = \sum_{r=0}^{k} a_r\, j^r \tag{7}$$

In order to solve for the coefficients of the polynomial, the principle of the least squares method is applied to minimize the sum of squared residuals between the fitted values and the actual observed values within the window.

$$E = \sum_{j=-m}^{m} \left(\sum_{r=0}^{k} a_r\, j^r - x_{i+j}\right)^2 \tag{8}$$

In order to minimize the error function, the partial derivatives with respect to each coefficient are taken and set to zero, allowing the coefficient matrix to be solved. In practical applications, through complex matrix operations and derivations, the smoothed fitting value at the center point is actually equivalent to the linearly weighted combination of the data points within the window. Therefore, the final smoothed output of the S-G filter can be expressed in the form of a convolution sum.

$$x_i^* = \sum_{j=-m}^{m} C_j\, x_{i+j} \tag{9}$$

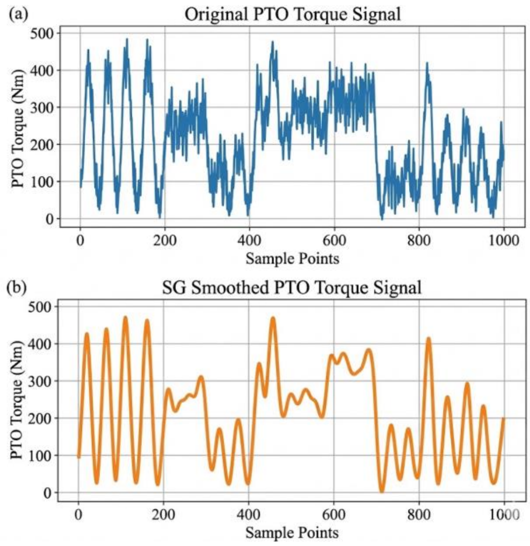

Fig. 3 – Comparison chart before and after noise reduction treatment

RESULTS

Load Counting Statistical Processing

After preprocessing, the data can be further screened and streamlined to improve analysis efficiency. The rain-flow counting method is a widely used statistical counting technique for measured load time histories, which extracts information on the amplitude and mean of load cycles, aligning with the inherent characteristics of fatigue loads. For the preprocessed load signal, the four-point rain-flow counting method is employed to perform statistical counting processing on the measured load, resulting in a mean-amplitude matrix of load cycles.

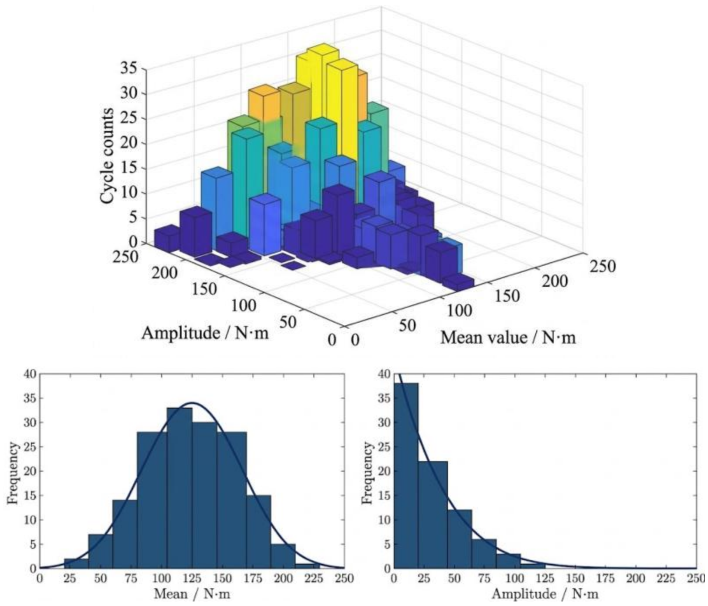

Fig. 4 – Schematic diagram of rainflow matrix

Kernel Density Extrapolation Based on DBSCAN Clustering

After completing the preprocessing of the tractor PTO load signal, the rain-flow counting method is used to extract load cycles from the torque time history, transforming the random load into a series of complete mean-amplitude cycles to form a sample set.

$$X = \{(m_i, a_i) \mid i = 1,2, \dots, N\} \qquad (10)$$

Among them, m_i is the mean value of the cycle, and a_i is the amplitude. This two-dimensional feature space serves as the foundation for subsequent clustering and probabilistic modeling. Due to the complex operating conditions of tractors, PTO loads often exhibit multi-modal distribution characteristics. Therefore, the DBSCAN density clustering algorithm is introduced to partition the sample set. DBSCAN identifies clusters of arbitrary shapes based on density reach-ability and automatically eliminates noise points.

The ε of a sample point x_i is defined as:

$$N_\varepsilon(x_i) = \{x_j \in X \mid d(x_i, x_j) \le \varepsilon\} \tag{11}$$

The Euclidean distance is usually used as the distance metric d(·). If the number of sample points within the neighborhood satisfies

$$|N_\varepsilon(x_i)| \ge MinPts \tag{12}$$

Then x_i is a core point. Clusters are continuously expanded through direct density-reachability and density-reachability relationships, ultimately partitioning the samples into K clusters C_1, C_2, …, C_K and several noise points. The selection of the parameters ε and $MinPts$ is crucial: the k-distance graph method is used to determine ε (taking k=$MinPts$), that is, the distance from each sample to its nearest neighbor is calculated and sorted in descending order, and the distance corresponding to the inflection point of the curve is taken as ε; $MinPts$ is empirically tuned between 4 and 10. The clustering algorithm is shown in the figure.

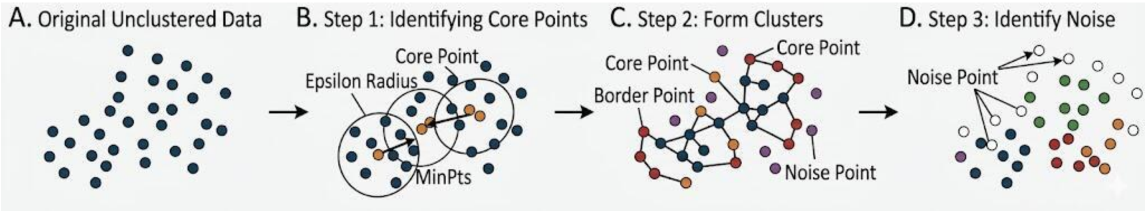

Fig. 5 – DBSCAN algorithm schematic diagram

After clustering is completed, a two-dimensional probability density function is established for each cluster C_j (with sample sizen_j), and kernel density estimation is used to describe its distribution. For the cluster, the probability density estimate is given by:

$$\hat{f}_j(m, a) = \frac{1}{n_j}\sum_{i=1}^{n_j} K_{H_j}(m - m_{ji}, a - a_{ji}) \tag{13}$$

where $K_{H_j}(\cdot)$ is the two-dimensional Gaussian kernel function.

$$K_{H_j}(u) = \frac{1}{2\pi|H_j|^{1/2}} exp\left(-\frac{1}{2}u^T H_j^{-1} u\right) \tag{14}$$

H_j is the bandwidth matrix. The selection of bandwidth directly affects the estimation accuracy, and the traditional global fixed-bandwidth struggles to accommodate the differences in data distribution across clusters. Therefore, the optimal bandwidth is calculated separately based on the characteristics of each cluster. The Rule-of-Thumb method is used to determine the bandwidth matrix. For two-dimensional data, n rule can be adopted:

$$.H_j = n_j^{-1/6}\hat{\Sigma}_j^{1/2} \tag{15}$$

where $\hat{\Sigma}_j$ is the sample covariance matrix of the cluster, estimated by the following formula:

$$\hat{\Sigma}_j = \frac{1}{n_j - 1}\sum_{i=1}^{n_j}\begin{bmatrix} (m_{ji} - \bar{m}_j)^2 & (m_{ji} - \bar{m}_j)(a_{ji} - \bar{a}_j) \\ (m_{ji} - \bar{m}_j)(a_{ji} - \bar{a}_j) & (a_{ji} - \bar{a}_j)^2 \end{bmatrix} \tag{16}$$

In the formula, $\bar{m}_j$ and $\bar{a}_j$ are the sample means of the mean and amplitude for that cluster, respectively. This adaptive bandwidth method preserves details in data-dense regions while providing smooth estimation in sparse regions, thereby improving the goodness of fit. After obtaining the probability density functions for each cluster, Monte Carlo simulation is used for load extrapolation. First, the sampling weights for each cluster are calculated:

$$w_j = \frac{n_j}{\sum_{j=1}^{K} n_j} \tag{17}$$

Noise points are not involved in modeling. The extrapolation factor is determined based on the operating time T_0 corresponding to the measured data and the design life T_{target}:

$$\lambda = \frac{T_{target}}{T_0} \tag{18}$$

The total number of extrapolated cycles is approximately:

$$N_{target} = \lambda N \tag{19}$$

The sampling process is as follows: a cluster is randomly selected according to the weight w_{j}j, and then a sample point is drawn from the kernel density mixture distribution of that cluster.

Specifically, a sample (m_{ji},a_{ji}) within the cluster is first randomly selected as the kernel center, and then a new sample (m_{newi},a_{newi}) is drawn from a two-dimensional Gaussian distribution with the center as the mean and a covariance matrix of H_{j2}. This process is repeated until N_{target} samples are obtained. To ensure the physical rationality of the extrapolated loads, boundary constraints are applied to the sampling results. Samples that exceed the boundaries are discarded and resampled. Finally, the extrapolated load cycle set is obtained:

$$X_{ext} = \{(m_k, a_k) \mid k = 1,2,\dots,N_{target}\} \tag{20}$$

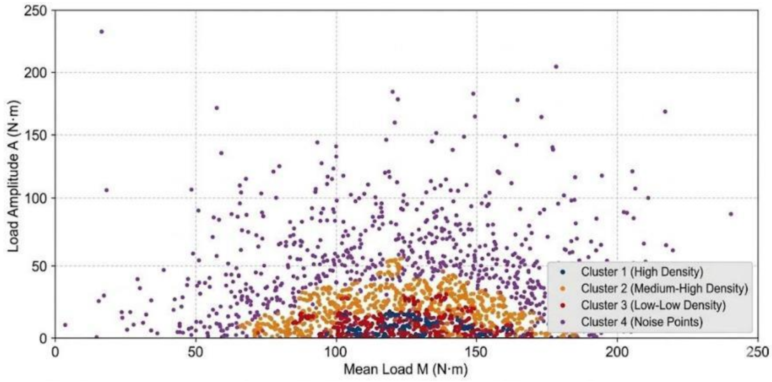

Fig. 6 – DBSCAN clustering effect diagram

Figure 6 shows the identification of load distribution based on DBSCAN clustering. Combined with adaptive bandwidth kernel density estimation and Monte Carlo sampling, it enables high-precision non-parametric extrapolation of complex tractor PTO loads.

Extrapolated Load Verification

To verify the effectiveness and accuracy of the proposed method, a comparison is made with the traditional load extrapolation based on fixed-bandwidth kernel density estimation. The evaluation is conducted in terms of two aspects, namely goodness of fit and pseudo-damage.

The fixed-bandwidth method and the DBSCAN clustering method proposed in this paper were both used to extrapolate to five times the original load. The coefficient of determination and mean square error were introduced as two parameters to compare the extrapolation effects, and the results are shown in Table1. The results indicate that, compared with the fixed-bandwidth method, the coefficient of determination of the proposed method increased by 3.401%, while the mean square error decreased by 39.154%. It can be concluded that the clustering-based load extrapolation with clustered data significantly improves the extrapolation performance, verifying the feasibility and effectiveness of the proposed method.

Table 1

Comparison of Extrapolation Effects

Extrapolation model	Coefficient of determination	Mean square error
Fixed-bandwidth	0.941	8.745
DBSCAN clustering	0.973	5.321

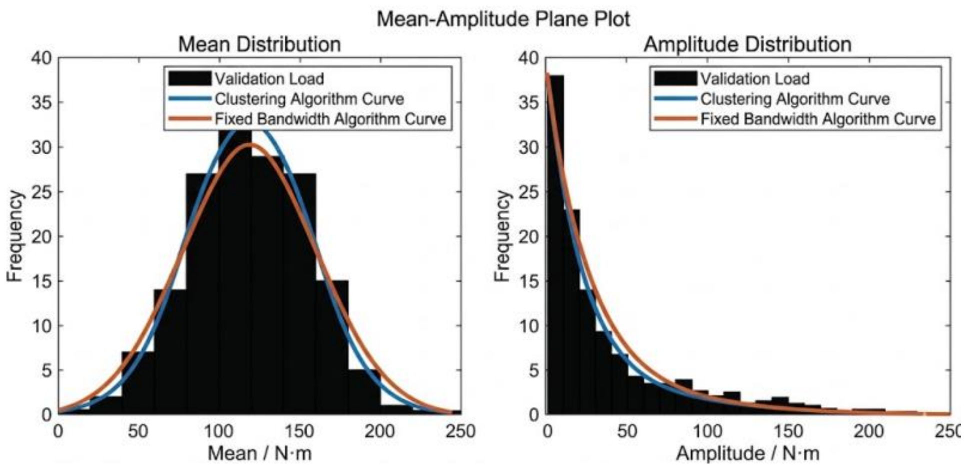

Fig. 7 – Mean amplitude probability fitting curve

To verify the accuracy of the extrapolated load spectrum generated by the DBSCAN based non-parametric estimation method, this paper introduces the pseudo-damage theory for equivalence testing. Pseudo-damage is an engineering indicator used to evaluate the severity of a load spectrum from the perspective of fatigue damage, comprehensively reflecting the influence of load amplitude distribution and cycle count on the fatigue life of components. The calculation of pseudo-damage is based on Miner's linear cumulative damage criterion and the S-N curve.

$$D_p = \sum_{i=1}^{k} n_i \, S_{ai}^{m} \qquad (21)$$

In the formula, D_p represents the total pseudo-damage of the load spectrum; k is the total number of load levels after rain flow counting and two-dimensional grid division of the load spectrum; S_{ai}^{m} is the torque amplitude of the load level; n_i is the cycle count corresponding to the load level; and m is taken as 6.

The pseudo-damage ratio R_d is defined as the ratio of the pseudo-damage of the extrapolated load spectrum to the pseudo-damage of the proportionally scaled original load spectrum.

$$R_d = \frac{D_{p,ext}}{D_{p,ori} \times E_f} \qquad (22)$$

In the formula, R_d is the pseudo-damage ratio, $D_{p,ext}$ is the pseudo-damage value of the full-life load spectrum generated by two-dimensional non-parametric extrapolation, $D_{p,ori}$ is the pseudo-damage value of the original measured load, and E_f is the extrapolation factor.

Table 2

Pseudo-damage ratio at different multiples

Different multiples	Fixed-bandwidth	The proposed method
10	1.098	1.032
20	1.112	1.038
30	1.103	1.041
40	1.113	1.043
50	1.121	1.045

Based on the comparison data of pseudo-damage ratios under different extrapolation factors, the adaptive bandwidth method based on DBSCAN clustering proposed in this paper demonstrates significantly superior performance compared to the traditional fixed-bandwidth method at all extrapolation factors. The pseudo-damage ratio of the fixed-bandwidth method fluctuates between 1.098 and 1.121, showing a clear upward trend as the extrapolation factor increases. In contrast, the pseudo-damage ratio of the proposed method remains stable between 1.032 and 1.045, not only closer to the ideal value of 1, but also exhibiting an increase of only 1.260% as the extrapolation factor rises, demonstrating excellent accuracy retention and factor insensitivity.

From a quantitative perspective, the method proposed in this paper reduces the deviation of the pseudo-damage ratio by an average of approximately 6.512% compared to the fixed-bandwidth method. The fundamental reason for this difference lies in the fact that the fixed-bandwidth kernel density estimation adopts a globally uniform smoothing parameter, which makes it difficult to simultaneously achieve fitting accuracy in both dense and sparse regions when dealing with the multi-modal distribution characteristics of tractor PTO loads, leading to distortion of the distribution tail features that is amplified with increasing extrapolation factors. In contrast, the DBSCAN clustering method effectively identifies natural clusters corresponding to different operating conditions through density-based clustering, and computes adaptive bandwidths based on the sample covariance of each cluster, thereby more accurately preserving the original load distribution morphology and fatigue damage characteristics.

The results indicate that the proposed method can more accurately replicate the fatigue damage potential of loads under different extrapolation requirements, exhibiting higher accuracy and robustness, and is suitable for load spectrum compilation throughout the full life cycle of tractor PTO.

ACKNOWLEDGMENTS

This study was funded by the key project "Collection and Compilation of High Power Tractor Operating Load Spectrum"(number: HTSP-20251217004) of Shaanxi FAST Auto Drive Group Co., Ltd.

CONCLUSIONS

(1) The clustering effectiveness of the DBSCAN algorithm is highly dependent on the selection of the neighborhood radius ε and the minimum number of samples $MinPts$. The current approach based on the k-distance graph combined with empirical tuning has a certain degree of subjectivity.

Future research can explore automatic parameter optimization methods based on clustering validity indices or introduce adaptive DBSCAN variants to achieve full adaptability in the clustering process, further enhancing the robustness and generality of the method. In non-parametric rain-flow extrapolation based on kernel density estimation, the kernel bandwidth is the most critical factor affecting the extrapolation performance. More accurate and effective selection algorithms can be introduced for the choice of kernel bandwidth.

(2) This study only considered the load characteristics in the two dimensions of mean and amplitude, whereas actual fatigue damage is also influenced by factors such as load sequence and frequency components. Subsequent research could attempt to incorporate additional features such as load duration and rise/fall rates into the analysis dimension to construct a higher-dimensional probability model. At the same time, methods for reconstructing time series with realistic temporal characteristics from extrapolated cycle distributions can be explored to meet the requirements of bench tests that demand full time-domain signals.

(3) The method proposed in this paper was validated based on PTO load data from a specific tractor model. Future work should extend the validation to load data from more tractor models and a wider range of operation types (such as plowing, rotary tillage, transportation, etc.) to verify the scalability of the method. More importantly, indoor bench durability tests using the extrapolated load spectrum should be conducted. By comparing the actual fatigue life of components with the life predicted by the extrapolated spectrum, a quantitative evaluation system for method accuracy can be established, providing a more solid foundation for engineering applications.

REFERENCES

[1] Anonymous (2018). Developing an accelerated flight load spectrum based on the n(z)-N curves of a fleet. *Journal of Technology & Science*, pp. 986, United States.

[2] Bučar, T., Nagode, M., Fajdiga, M. (2003). Reliability approximation using finite Weibull mixture distributions. *Reliability Engineering and System Safety*, Vol. 84, pp. 241-251, Netherlands.

[3] Chen, C., Yang, Z., He, J., (2017). Load spectrum generation of machining center based on rainflow counting method. *Journal of Vibroengineering*, Vol. 19, pp. 5767-5779, Lithuania.

[4] Chen, D., Xiang, C., Chen, X. (2002). Load spectrum establishment for power train parts of military vehicle. *Journal of Mechanical Strength*, Vol. (02), pp. 310-314, China.

[5] He, J., Wang, S., Li, G., (2018). Compilation of NC lathe dynamic cutting force spectrum based on two-dimensional mixture models. *The International Journal of Advanced Manufacturing Technology*, Vol. 98, pp. 251-262, United Kingdom.

[6] Jin, X., Zhang, S. (2019). Nonparametric fitting extrapolation based on measured load of high-speed axle. *Journal of Beijing Jiaotong University*, Vol. 43, pp. 97-103+130, China.

[7] Liu, Y., Zhang, X., Wang, Z., (2011). Comparison of load spectrum extrapolation methods. *Modern Manufacturing Engineering*, Vol. (11), pp. 8-11, China.

[8] Liu, Y., Chang, L., Sun, L. (2012). Research on load spectrum compilation method for wheel loader driveline. *China Mechanical Engineering*, Vol. 23, pp. 1412-1416, China.

[9] Li, Y., Zhao, P., Gu, Z., Cheng, L., Li, S. (2025). Transformer-Based Load Spectrum Compilation for Rolling Bearings. *Journal of Engineering*, Vol. 2025, pp. 5876690-5876690, United Kingdom.

[10] Li, S., Zhai, Y., Wang, X., (2021). Load extrapolation method for tractor drive shaft based on kernel density estimation. *Journal of China Agricultural University*, Vol. 26, pp. 175-184, China.

[11] Li, Y. (2016). Research on nonparametric extrapolation method in load spectrum compilation of wheel loader [D]. Jilin University, China.

[12] Liu, H., Zhang, W., Xu, Y., (2017). Load spectrum compilation of tracked vehicle drive shaft based on kernel density estimation. *Acta Armamentarii*, Vol. 38, pp. 1830-1838, China.

[13] Nagode, M., Klemenc, J., Fajdiga, M. (2001). Parametric modelling and scatter prediction of rainflow matrices. *International Journal of Fatigue*, Vol. 23, pp. 525-532, Netherlands.

[14] Tian, W., Song, X., Wan, Y. (2019). Overview of common load spectrum extrapolation algorithms. *Southern Agricultural Machinery*, Vol. 50, pp. 103+108, China.

[15] Wang, L., Zong, J., Wang, Y., (2022). Compilation of Tractor Three-Point Hitch Traction Load Spectrum and Bench Test Based on Optimal Distribution Fitting. *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 38, pp. 41-49, China.
[16] Wang, J., Chen, X., (2016). A Review of the Extrapolation Method in Load Spectrum Compiling. *Journal of Mechanical Engineering*, Vol. 62, pp. 60-75, China.
[17] Wang, J., Zhang, X., Liu, C., (2017). Determination of the Threshold for Extreme Load Extrapolation Based on Multi-Criteria Decision-Making Technology. *Strojniški vestnik - Journal of Mechanical Engineering*, Vol. 63, pp. 201-211, Slovenia.
[18] Wang, J., Zhang, Z., Luo, L., (2021). SwinGD: A robust grape bunch detection model based on Swin transformer in complex vineyard environments. *Horticulturae*, Vol. 7, pp. 492, Switzerland.
[19] Yang, Z., Song, Z., Luo, Z., (2022). Time domain extrapolation method for key component loads of tractor based on EMD-POT model. *Journal of Mechanical Engineering*, Vol. 58, pp. 252-262, China.
[20] Yu, L., An, Y., He, J., (2023). Research progress and development trend of load spectrum extrapolation technology for electromechanical equipment. *Journal of Jilin University (Engineering and Technology Edition)*, Vol. 53, pp. 941-953, China.
[21] Zhang, T., Wu, F., Wang, M., (2023). Grape-bunch identification and picking point location on occluded fruit axes based on YOLOv5-GAP. *Horticulturae*, Vol. 9, pp. 498, Switzerland.